Long time behavior of diffusions with Markov switching

Jean-Baptiste Bardet, Hélène Guérin, Florent Malrieu

Preprint – December 17, 2009

Abstract

Let Y be an Ornstein-Uhlenbeck diffusion governed by an ergodic finite state Markov process X: $dY_t = -\lambda(X_t)Y_t dt + \sigma(X_t)dB_t$, Y_0 given. Under ergodicity condition, we get quantitative estimates for the long time behavior of Y. We also establish a trichotomy for the tail of the stationary distribution of Y: it can be heavy (only some moments are finite), exponential-like (only some exponential moments are finite) or Gaussian-like (its Laplace transform is bounded below and above by Gaussian ones). The critical moments are characterized by the parameters of the model.

AMS Classification 2000: 60J60, 60J75, 60H25.
Key words: Ornstein-Uhlenbeck diffusion, Markov switching, jump process, random difference equation, light tail, heavy tail, Laplace transform, convergence to equilibrium.

1 Introduction and main results

The aim of this paper is to draw a complete picture of the ergodicity of Ornstein-Uhlenbeck diffusions with Markov switching (characterization of the tails of the invariant measure and quantitative convergence to equilibrium). In particular we make more precise the results of [7, 4]. The so-called diffusion with Markov switching $Y = (Y_t)_{t\geqslant 0}$ is defined as follows.

The *switching process* $X = (X_t)_{t\geqslant 0}$ is a Markov process on the finite state space $E = \{1, \ldots, d\}$ (with $d \geqslant 2$), of infinitesimal generator $A = (A(x, \tilde{x}))_{x,\tilde{x}\in E}$. Let us denote by $a(x)$ the jump rate at state $x \in E$ and $P = (P(x, \tilde{x}))_{x,\tilde{x}\in E}$ the transition matrix of the embedded chain. One has, for $x \neq \tilde{x}$ in E,

$$a(x) = -A(x, x) \quad \text{and} \quad P(x, \tilde{x}) = -\frac{A(x, \tilde{x})}{A(x, x)}.$$

We assume that P is irreducible recurrent. The process X is ergodic with a unique invariant probability measure denoted by μ. See [10] for details. Let $\mathcal{F}_t^X = \sigma(X_u, 0 \leqslant u \leqslant t)$. Moreover, let $\mathbb{E}_x$ denote the expectation with respect to the law $\mathbb{P}_x$ of X knowing that $X_0 = x$.

Let $B = (B_t)_{t\geqslant 0}$ be a standard Brownian motion on $\mathbb{R}$ and Y_0 a real-valued random variable such that B, Y_0 and X are independent. Conditionnally to X, the process $Y = (Y_t)_{t\geqslant 0}$ is the real-valued diffusion process defined by:

$$Y_t = Y_0 - \int_0^t \lambda(X_u)Y_u \, du + \int_0^t \sigma(X_u) \, dB_u, \tag{1}$$

where λ and σ are two functions from E to $\mathbb{R}$ and $(0,\infty)$ respectively. Of course, if λ and σ are constant, Y is just an Ornstein-Uhlenbeck process with attractive ($\lambda > 0$), neutral ($\lambda = 0$) or repulsive coefficient ($\lambda < 0$). One has to notice that Equation (1) has an "explicit" solution:

$$Y_t = Y_0 \exp\left(-\int_0^t \lambda(X_u)\,du\right) + \int_0^t \exp\left(-\int_u^t \lambda(X_v)\,dv\right)\sigma(X_u)\,dB_u. \tag{2}$$

Remark 1.1. *In others words, the full process (X,Y) is the Markov process on $E\times\mathbb{R}$ associated to the infinitesimal generator $\mathcal{A}$ defined by:*

$$\mathcal{A}f(x,y) = \sum_{\tilde{x}\in E} A(x,\tilde{x})(f(\tilde{x},y) - f(x,y)) + \frac{\sigma(x)^2}{2}\partial^2_{22} f(x,y) - \lambda(x)\partial_2 f(x,y).$$

Previous works investigated the ergodicity of Y and some integrability properties for the invariant measure. For example, in [2], the multidimensional case is adressed together with the case of diffusion coefficients depending on Y. Stability results and sufficient conditions for the existence of moments are established under Lyapunov-type conditions.

In [7], it is proved that the Markov switching diffusion Y is ergodic if and only if

$$\sum_{x\in E} \lambda(x)\mu(x) > 0, \tag{3}$$

that is if the process is attractive "in average". Let us denote by ν its invariant probability measure of Y. It is also shown in [7] that ν admits a moment of order p if, for any $x\in E$, $p\lambda(x) + a(x)$ is positive and the spectral radius of the matrix

$$M_p = \left(\frac{a(x)}{a(x)+p\lambda(x)} P(x,\tilde{x})\right)_{x,\tilde{x}\in E} \tag{4}$$

is smaller than 1. In the sequel $\rho(M)$ stands for the spectral radius of a matrix M.

In [4], the result is more precise: a dichotomy is exhibited between heavy and light tails for ν. Let us define

$$\underline{\lambda} = \min_{x\in E} \lambda(x) \quad \text{and} \quad \overline{\lambda} = \max_{x\in E} \lambda(x). \tag{5}$$

Theorem 1.2 (de Saporta-Yao [4])**.** *Under Assumption* (3)*, the following dichotomy holds:*

1. *if $\underline{\lambda} < 0$, then there exists $C > 0$ such that*

$$t^\kappa \nu((t,+\infty)) \xrightarrow[t\to+\infty]{} C,$$

where κ is the unique $p \in (0, \min\{-a(x)/\lambda(x),\ \lambda(x) < 0\})$ such that the spectral radius of M_p is equal to 1;

2. *if $\underline{\lambda} \geqslant 0$, then ν has moments of all order.*

Remark 1.3. *Note that the constant κ does not depend on the parameters $(\sigma(x))_{x\in E}$, and that Point 1. from previous theorem implies that, for $\underline{\lambda} < 0$, the p^{th} moment of ν is finite if and only if $p < \kappa$.*
The main idea of the proofs in [7] and [4] is to study the discrete time Markov chain $(X_{\delta n}, Y_{\delta n})_{n\geqslant 0}$ for any $\delta > 0$ with renewal theory and then to let δ goes to 0.

The main goal of the present paper is to show that there are three (and not only two) different behaviors for the tails of ν.

Let us gather below several useful notations.

Notations 1.4. *Let us define for the diffusion coefficients*

$$\underline{\sigma}^2 = \min_{x\in E} \sigma^2(x) \quad \text{and} \quad \overline{\sigma}^2 = \max_{x\in E} \sigma^2(x). \tag{6}$$

We denote by A_p the matrix $A - p\Lambda$ where Λ is the diagonal matrix with diagonal $(\lambda(1), \dots, \lambda(d))$ and associate to A_p the quantity

$$\eta_p := - \max_{\gamma\in \mathrm{Spec}(A_p)} \mathrm{Re}\ \gamma. \tag{7}$$

When $\underline{\lambda} \geqslant 0$, the set E is the union of

$$M = \{x \in E,\ \lambda(x) > 0\} \quad \text{and} \quad N = \{x \in E,\ \lambda(x) = 0\}. \tag{8}$$

Let us then define

$$\beta(x) = \frac{\sigma(x)^2}{2a(x)} \quad \text{and} \quad \overline{\beta} = \max_{x\in N} \beta(x), \tag{9}$$

and, for any v such that $v^2 < \overline{\beta}^{-1}$, the matrix

$$P_v^{(N)} = \left(\frac{1}{1-\beta(x)v^2} P(x,x')\right)_{x,x'\in N}. \tag{10}$$

We are now able to state our main result.

Theorem 1.5. *Let us define*

$$\kappa = \sup\{p \geqslant 0,\ \eta_p > 0\} \in (0, +\infty].$$

Then η_p is continuous, positive on the set $(0, \kappa)$ and negative on $(\kappa, +\infty)$. Under Assumption (3), the following trichotomy holds:

1. *if $\underline{\lambda} < 0$ then $0 < \kappa \leqslant \min\{-a(x)/\lambda(x),\ \lambda(x) < 0\}$, and the p^{th} moment of ν is finite if and only if $p < \kappa$.*

2. *if $\underline{\lambda} = 0$, then κ is infinite and the domain of the Laplace transform of ν is $(-v_c, v_c)$ where*
$$v_c = \sup\left\{v > 0,\ \rho(P_v^{(N)}) < 1\right\}; \tag{11}$$

3. *if $\underline{\lambda} > 0$, then κ is infinite and ν has a Gaussian-like Laplace transform: for any $v \in \mathbb{R}$,*
$$\exp\left(\frac{\underline{\sigma}^2 v^2}{4\overline{\lambda}}\right) \leqslant \int e^{vy}\,\nu(dy) \leqslant \exp\left(\frac{\overline{\sigma}^2 v^2}{4\underline{\lambda}}\right).$$
Moreover, its tail looks like the one of the Gaussian law with variance $\overline{\alpha}/2$ where $\overline{\alpha} = \max_{x\in E} \sigma(x)^2/\lambda(x)$ since $y \mapsto e^{\delta y^2}$ is ν-integrable if and only if $\delta < 1/\overline{\alpha}$.

Remark 1.6. *In the sequel we will respectively refer to Points 1. 2. and 3. as the polynomial, exponential-like and Gaussian-like cases.*

The first point of this theorem is a reformulation of the first point of Theorem 1.2 by de Saporta and Yao. We can in particular check that our characterization of κ in Theorem 1.5 is equivalent to the one given by de Saporta and Yao in Point 1. of Theorem 1.2 (see Remark 4.3). We provide a direct and simple proof of this result based on Itô formula and some basic results on finite Markov chains. The proof of Points 2. relies on precise estimates on the Laplace transform of Y_t that can be derived from a discrete time model already studied in [6, 8, 1].

It is straightforward from (2) that, for any measure π_0 on $E\times\mathbb{R}$, the Laplace transform L_t of Y_t is
$$L_t(v) := \mathbb{E}_{\pi_0}\left(e^{vY_t}\right) = \mathbb{E}_{\pi_0}\left[\exp\left(vY_0 e^{-\int_0^t \lambda(X_s)\,ds} + \frac{v^2}{2}\int_0^t \sigma(X_s)^2 e^{-2\int_s^t \lambda(X_r)\,dr}\,ds\right)\right]. \quad (12)$$

The estimate of the Laplace transform in the Gaussian-like case (Point 3.) is hence easily deduced from this explicit expression. Assuming that $Y_0 = 0$, we get from (12) that
$$L_t(v) \leqslant \mathbb{E}\left[\exp\left(\frac{v^2}{2}\int_0^t \overline{\sigma}^2 e^{-2\int_s^t \underline{\lambda}\,dr}\,ds\right)\right] \leqslant \exp\left(\left(1-e^{-2\underline{\lambda}t}\right)\frac{\overline{\sigma}^2 v^2}{4\underline{\lambda}}\right),$$
which gives the upper bound as t goes to infinity. The lower bound follows from a symmetric argument.

The proofs of Point 2. and of the second part of Point 3. are more delicate (and interesting). For the exponential case, we first get the critical exponential moment for the process Y observed at the hitting times of the subset M defined in (8). Then we show that the full process has the same critical exponent.

At the end of the paper we focus on the convergence of the law of Y_t to the invariant measure ν. We get an explicit exponential bound for the Wasserstein distance of order p for any $p < \kappa$. Classically, let $p \geqslant 1$ and $\mathcal{P}_p$ be the set of the probability measures on $\mathbb{R}$ with a finite p^{th} moment. Define the Wasserstein distance W_p on $\mathcal{P}_p$ as follows: for any ρ and $\tilde{\rho}$ in $\mathcal{P}_p$,
$$W_p(\rho,\tilde{\rho}) = \left(\inf_{\pi}\left\{\int |y-\tilde{y}|^p\,\pi(dy,d\tilde{y})\right\}\right)^{1/p},$$
where the infimimum is taken among all the probability measures π on $\mathbb{R}^2$ with marginals ρ and $\tilde{\rho}$. It is well-known that $(\mathcal{P}_p, W_p)$ is a complete metric space (see [11]).

The strategy is to couple two processes (X,Y) and $(\tilde{X},\tilde{Y})$ in such a way that the Wasserstein distance between $\mathcal{L}(Y_t)$ and $\mathcal{L}(\tilde{Y}_t)$ goes to zero as t goes to infinity. This requires to couple the initial conditions and the dynamics (of both the Markov chains and the diffusion part). When X_0 and $\tilde{X}_0$ have the same law, the coupling is trivial: we choose $X = \tilde{X}$ and the same driving Brownian motion.

Theorem 1.7. *Let $p < \kappa$. Assume that X_0 and $\tilde{X}_0$ have the same law. Let Y and $\tilde{Y}$ be solutions of* (1) *associated to (X_t) and $(\tilde{X}_t)$ and assume that Y_0 and $\tilde{Y}_0$ have finite moment of order p. Then there exists $C(p)$ such that*

$$W_p\Big(\mathcal{L}(Y_t), \mathcal{L}(\tilde{Y}_t)\Big)^p \leqslant C(p) e^{-\eta_p t} W_p\Big(\mathcal{L}(Y_0), \mathcal{L}(\tilde{Y}_0)\Big)^p,$$

where η_p is given by (7).

If X_0 and $\tilde{X}_0$ do not have the same law, one first has to make the Markov chains X and $\tilde{X}$ stick together and then to use Theorem 1.7. This provides a rather intricate bound which is given for convenience in Section 5.

The paper is organised as follows. In Section 2 we complete the proof for the Gaussian-like case of Theorem 1.5. The exponential-like case is studied in Section 3. Since the critical exponential moment is not explicit in the general case, we give also the explicit computation of the Laplace transform of ν when E is reduced to $\{1, 2\}$. In Section 4 we establish a uniform bound for the p^{th} moment of $(Y_t)_t$ for any $p < \kappa$ and the first point of Theorem 1.5 as a corollary. We finally provide the proof of Theorem 1.7 and its extension to general initial conditions in Section 5.

2 Gaussian moments for the switched diffusion

This section is dedicated to the proof of the second part of Point 3. of Theorem 1.5.

Proof of Point 3. of Theorem 1.5. Let us denote by

$$\alpha(x) = \frac{\sigma(x)^2}{\lambda(x)} \quad \text{for } x \in E \text{ and } \quad \overline{\alpha} = \max_{x \in E} \alpha(x) < +\infty.$$

For any $\delta \in (0, 1/\overline{\alpha})$, Itô's formula ensures that

$$de^{\delta Y_t^2} = \big(-2\lambda(X_t)\delta Y_t^2 + (2\delta^2 Y_t^2 + \delta)\sigma(X_t)^2\big) e^{\delta Y_t^2} dt + dM_t$$

where $(M_t)_t$ is a martingale. For any $x \in E$ and $y \in \mathbb{R}$,

$$\begin{aligned} 2(-\lambda(x) + \delta\sigma(x)^2) y^2 + \sigma(x)^2 &\leqslant -2\lambda(x)(1 - \delta\overline{\alpha}) y^2 + \overline{\alpha}\lambda(x) \\ &\leqslant -2\underline{\lambda}(1 - \delta\overline{\alpha}) y^2 + \overline{\alpha}\overline{\lambda}, \end{aligned}$$

since $\delta\overline{\alpha} < 1$. Moreover, for any $a > 0$, there exists $b > 0$ such that, for any $y \in \mathbb{R}$,

$$-2\underline{\lambda}\delta(1 - \overline{\alpha}\delta) y_t^2 + \overline{\lambda}\overline{\alpha}\delta \leqslant -a + be^{-\delta y^2},$$

thus

$$\frac{d}{dt}\mathbb{E}\Big(e^{\delta Y_t^2}\Big) \leqslant -a\mathbb{E}\Big(e^{\delta Y_t^2}\Big) + b.$$

As a consequence, $\sup_{t \geqslant 0} \mathbb{E}\Big(e^{\delta Y_t^2}\Big)$ is finite as soon as $\mathbb{E}\Big(e^{\delta Y_0^2}\Big)$ is finite and $\delta\overline{\alpha} < 1$.

On the other hand, assume (without loss of generality) that $\alpha(1) = \overline{\alpha}$. Choose (X_0, Y_0) with law $\overline{\nu}$ (the invariant measure of (X, Y)). For any $t > 0$, we have

$$\mathbb{E}\Big(e^{\delta Y_0^2}\Big) = \mathbb{E}\Big(e^{\delta Y_t^2}\Big) \geqslant \mathbb{E}\Big[\mathbb{1}_{\{X_0 = 1\}} \mathbb{E}_{1, Y_0}\Big(\mathbb{1}_{\{T_1 > t\}} e^{\delta Y_t^2}\Big)\Big],$$

where T_1 is the first jump time of X. On the set $\{T_1 > t\}$,

$$Y_t \overset{\mathcal{L}}{=} Y_0 e^{-\lambda(1)t} + N_t$$

where N_t is a centered Gaussian random variable with variance $\alpha(1)(1-e^{-2\lambda(1)t})/2$ which is independent of Y_0 and T_1. Thus, reminding that $T_1 \sim \mathcal{E}(a(1))$, we get

$$\mathbb{E}_{1,Y_0}\Big(\mathbb{1}_{\{T_1>t\}} e^{\delta Y_t^2}\Big) = e^{-a(1)t}\mathbb{E}\Big(e^{\delta(Y_0 e^{-\lambda(1)t}+N_t)^2}\Big).$$

Since $a \mapsto \mathbb{E}\Big(e^{\delta(a+N_t)^2}\Big)$ is even and convex, it reaches its minimum at $a=0$ and

$$\mathbb{E}\Big(e^{\delta(Y_0 e^{-\lambda(1)t}+N_t)^2}\Big) \geqslant \mathbb{E}\Big(e^{\delta N_t^2}\Big) = \begin{cases} \dfrac{1}{\sqrt{1-\delta\alpha(1)(1-e^{-2\lambda(1)t})}} & \text{if } \delta\alpha(1)(1-e^{-2\lambda(1)t}) < 1, \\ +\infty & \text{otherwise.} \end{cases}$$

As a consequence, if $\delta > 1/\alpha(1)$, $\mathbb{E}\Big(e^{\delta Y_t^2}\Big)$ is bounded below by a function of t which is infinite for t large enough. Thus, $\mathbb{E}\Big(e^{\delta Y_t^2}\Big)$ is infinite too. $\square$

3 Exponential moments for the switched diffusion

This section is dedicated to the proof of Point 2. in Theorem 1.5. We assume in the sequel that $\underline{\lambda} = 0$. If $(X_t)_{t\geqslant 0}$ is a two-states Markov process then one can use (12) to compute explicitely the Laplace transform of the invariant measure ν. This is a warm-up for the general case, and gives a more explicit formula for the critical exponential moment, whereas it will come from an abstract spectral criterion in the general case.

3.1 The explicit expression for the two-states case

In this subsection we assume that $E = \{1,2\}$ and that $\underline{\lambda} = 0$. Let us start with a straightforward computation which suggests that the Laplace transform of the invariant measure of Y is infinite outside a bounded interval.

Remark 3.1. *If T is an exponential random variable with parameter a and B is a standard Brownian motion on $\mathbb{R}$ (with T and B independent) then,*

$$\mathbb{E}\big(e^{v\sigma B_T}\big) = \int_0^\infty \mathbb{E}\big(e^{v\sigma B_t}\big) a e^{-at}\,dt = \int_0^\infty e^{\sigma^2 v^2 t/2} a e^{-at}\,dt = \frac{2a}{2a-\sigma^2 v^2}.$$

In other words, the law of σB_T is a (symmetric) Laplace law. When X spends an exponential time in $x \in E$ with $\lambda(x) = 0$, Y behaves like $\sigma(x)B$.

Theorem 3.2 (The two-states degenerate case)**.** *Assume that $E = \{1,2\}$, $\lambda(1) = \lambda > 0$ and $\lambda(2) = 0$. Then, for any v such that $v^2 < 1/\beta(2)$ (see (9) for the definition of β),*

$$L(v) = \int_{-\infty}^{+\infty} e^{vx}\nu(dx) = \left(\frac{1-\mu(1)\beta(2)v^2}{1-\beta(2)v^2}\right)\left(\frac{1}{1-\beta(2)v^2}\right)^{1+a(1)/\lambda} \exp\left(\frac{\sigma(1)^2 v^2}{4\lambda}\right). \quad (13)$$

If $v^2 \geqslant 1/\beta(2)$, $L(v)$ is infinite.

Proof. Since $E = \{1,2\}$, X is symmetric with respect to μ which is given by $\mu(1) = a(2)/(a(1)+a(2))$. Let us denote by L_t the Laplace transform of Y_t when $Y_0 = 0$ and X is stationnary *i.e.* $\mathcal{L}(X_0) = \mu$. From Equation (12), one has for any $v \in \mathbb{R}$,

$$\begin{aligned} L_t(v) &= \mathbb{E}_\mu\left[\exp\left(\frac{v^2}{2}\int_0^t \sigma(X_s)^2 e^{-2\int_s^t \lambda(X_r)\,dr}\,ds\right)\right] \\ &= \mathbb{E}_\mu\left[\exp\left(\frac{v^2}{2}\int_0^t \sigma(X_s)^2 e^{-2\int_0^s \lambda(X_r)\,dr}\,ds\right)\right] \end{aligned}$$

since μ is reversible. By monotone convergence, we get that, for any $v \in \mathbb{R}$,

$$L(v) = \mathbb{E}_\mu\left[\exp\left(\frac{v^2}{2}\int_0^\infty \sigma(X_s)^2 e^{-2\int_0^s \lambda(X_r)\,dr}\,ds\right)\right] \in [1,+\infty],$$

where L is the Laplace transform of ν.

Let us introduce two auxilliary functions: for $x = 1,2$,

$$L_x(v) = \mathbb{E}_x\left[\exp\left(\frac{v^2}{2}\int_0^\infty \sigma(X_s)^2 e^{-2\int_0^s \lambda(X_r)\,dr}\,ds\right)\right].$$

It is clear that

$$L(v) = \mu(1)L_1(v) + \mu(2)L_2(v).$$

Moreover, if for any $t \geqslant 0$, $\mathcal{F}_t = \sigma(X_s,\ 0 \leqslant s \leqslant t)$ and T is the first jump time of X, then

$$\begin{aligned} L_x(v) &= \mathbb{E}_x\left[\mathbb{E}_x\left\{\exp\left(\frac{v^2}{2}\int_0^\infty \sigma(X_s)^2 e^{-2\int_0^s \lambda(X_r)\,dr}\,ds\right)\Big|\mathcal{F}_T\right\}\right] \\ &= \mathbb{E}_x\left[\exp\left(\frac{v^2}{2}\int_0^T \sigma(X_s)^2 e^{-2\int_0^s \lambda(X_r)\,dr}\,ds\right)E_{x,T}\right], \end{aligned}$$

where

$$E_{x,T} = \mathbb{E}_x\left\{\exp\left(\frac{v^2}{2}\int_T^\infty \sigma(X_s)^2 e^{-2\int_0^s \lambda(X_r)\,dr}\,ds\right)\Big|\mathcal{F}_T\right\}.$$

For any $s \in [0,T[$, $X_s = x$ and then

$$\int_0^T \sigma(X_s)^2 e^{-2\int_0^s \lambda(X_r)\,dr}\,ds = \sigma(x)^2\frac{1-e^{-2\lambda(x)T}}{2\lambda(x)},$$

with the convention $(1-e^{-0\times T})/0 = T$. Similarly, for $t \geqslant T$,

$$\int_T^\infty \sigma(X_s)^2 e^{-2\int_0^s \lambda(X_r)\,dr}\,ds = e^{-2\lambda(x)T}\int_T^\infty \sigma(X_s)^2 e^{-2\int_T^s \lambda(X_r)\,dr}\,ds$$

The Markov property implies

$$\mathbb{E}_x\left[\exp\left(\frac{v^2}{2}\int_T^\infty \sigma(X_s)^2 e^{-2\int_0^s \lambda(X_r)\,dr}\,ds\right)\Big|\mathcal{F}_T\right] = L_{X_T}\left(ve^{-\lambda(x)T}\right).$$

Thus,

$$L_x(v) = \mathbb{E}\left[\exp\left(\frac{v^2\sigma(x)^2(1-e^{-2\lambda(x)T})}{4\lambda(x)}\right)L_{3-x}\left(ve^{-\lambda(x)T}\right)\Big|X_0 = x\right].$$

More precisely,

$$L_1(v) = \mathbb{E}_1\Big[\exp\left(\frac{v^2\sigma(1)^2(1-e^{-2\lambda T})}{4\lambda}\right)L_2\Big(ve^{-\lambda T}\Big)\Big],$$

and

$$L_2(v) = \mathbb{E}_2\Big[e^{v^2\sigma(2)^2T/2}L_1(v)\Big] = \begin{cases} \dfrac{2a(2)}{2a(2)-\sigma(2)^2v^2}L_1(v) & \text{if } \sigma(2)^2v^2 < 2a(2), \\ +\infty & \text{otherwise.} \end{cases}$$

Using $\beta(2) = \sigma(2)^2/2a(2)$, one easily gets that L_1 satisfies the following equation: for any $v^2 < 1/\beta(2)$,

$$\begin{aligned} L_1(v) &= \frac{1}{1-\beta(2)v^2}\int_0^\infty \exp\left(\frac{\sigma(1)^2v^2(1-e^{-2\lambda t})}{4\lambda}\right)L_1(ve^{-\lambda t})a(1)e^{-a(1)t}\,dt \\ &= \frac{1}{1-\beta(2)v^2}\int_0^1 \exp\left(\frac{\sigma(1)^2v^2(1-u^2)}{4\lambda}\right)L_1(vu)\frac{a(1)}{\lambda}u^{a(1)/\lambda-1}\,du. \end{aligned}$$

With $x = uv$,

$$L_1(v) = \frac{1}{1-\beta(2)v^2}\left(\frac{1}{v}\right)^{a(1)/\lambda} e^{\sigma(1)^2v^2/(4\lambda)}\int_0^v e^{-\sigma(1)^2x^2/(4\lambda)}\frac{a(1)}{\lambda}x^{a(1)/\lambda-1}L_1(x)\,dx.$$

Deriving this relation provides

$$L_1'(v) = \left(\frac{\beta(2)v}{1-\beta(2)v^2} - \frac{a(1)}{\lambda v} + \frac{\sigma(1)^2v}{2\lambda} + \frac{1}{1-\beta(2)v^2}\frac{a(1)}{\lambda v}\right)L_1(v).$$

Then L_1 is solution of

$$L_1'(v) = \left(\frac{\sigma(1)^2v}{2\lambda} + \frac{\beta(2)(1+a(1)/\lambda)v}{1-\beta(2)v^2}\right)L_1(v)$$

which leads to

$$L_1(v) = e^{\sigma(1)^2v^2/(4\lambda)}\left(\frac{1}{1-\beta(2)v^2}\right)^{1+a(1)/\lambda},$$

since $L_1(0) = 1$. Since L_2 is a function of L_1 we get

$$L(v) = e^{\sigma(1)^2v^2/(4\lambda)}\left(\frac{1-\mu(1)\beta(2)v^2}{1-\beta(2)v^2}\right)\left(\frac{1}{1-\beta(2)v^2}\right)^{1+a(1)/\lambda}.$$

□

3.2 The exponential-like case

In this subsection we provide the proof of Point 2. ($\underline{\lambda} = 0$) of Theorem 1.5. We first recall that, in this case, we split the state space E of the switching process X in two subsets M and N defined in (8). We denote also by F the points of M that can be reached in one step from N:

$$F = \left\{ x \in M,\ \sum_{\tilde{x} \in N} P(\tilde{x}, x) > 0 \right\}.$$

Assume for simplicity that $X_0 \in M$ and define by induction the sequence of times $(T_n)_{n \geqslant 0}$ by $T_0 = 0$ and, for $n \geqslant 0$,

$$T_{2n+1} = \inf\{t > T_{2n},\ X_t \in N\}, \quad \text{and} \quad T_{2n+2} = \inf\{t > T_{2n+1},\ X_t \in M\}.$$

When X is in M, Y looks like a Ornstein-Uhlenbeck process (with variable but attractive drift) while it looks like a Brownian motion (with variable but bounded below and above variance) when X is in N. Thus, heuristically the process Y might be larger after a sojourn of X in N than in M.

Let us notice that for $x \in N$,

$$Y_T = Y_0 + I_x \quad \text{where} \quad I_x = \int_0^T \sigma(X_s^x)\, dB_s$$

and X^x is the process X starting at x and T is the first hitting time of M. Our strategy is to determine the domain of the Laplace transform of I_x and then to establish that is also the one of the process Y at the entrance times of X into the set M *i.e* at the times $(T_{2n})_{n \geqslant 0}$. We will then extend the result to the full process (X, Y).

Proposition 3.3. *Under previous assumptions, for any* $v^2 < \overline{\beta}^{-1}$*, the two following conditions are equivalent:*

1. *for any* $x \in N$, $\mathbb{E}(e^{vI_x}) < +\infty$*;*
2. $\rho(P_v^{(N)}) < 1$*, where* $P_v^{(N)}$ *is defined in Equation* (10).

Proof. Let $x_0, x_1, \ldots, x_{n-1}$ be in N. We denote by $(Z_n)_n$ the embedded chain of X. On the set $H = \{Z_0 = x_0, \ldots, Z_{n-1} = x_{n-1}, Z_n \in M\}$,

$$I_{x_0} = \int_0^T \sigma(X_s^{x_0})\, dB_s = \sum_{j=0}^{n-1} \sigma(x_j) \sqrt{\tau_{x_j}} G_j,$$

where the random variables $(G_j)_j$, $(\tau_{x_j})_j$ are independent and $\mathcal{L}(G_j) = \mathcal{N}(0,1)$ and $\mathcal{L}(\tau(x_j)) = \mathcal{E}(a(x_j))$. As a consequence,

$$\mathbb{E}\big(e^{vI_{x_0}} | H\big) = \prod_{j=0}^{n-1} \mathbb{E}\left[\exp\left(\frac{v^2 \sigma(x_j)^2}{2} \tau_{x_j}\right)\right] = \prod_{j=0}^{n-1} \frac{1}{1 - \beta(x_j) v^2}.$$

One just computes

$$\mathbb{E}(e^{vI_{x_0}}) = \sum_{\substack{n\geqslant 1\\ x_1,\dots,x_{n-1}\in N}} \mathbb{E}(e^{vI_{x_0}} \mid Z_1 = x_1, \dots, Z_{n-1} = x_{n-1}, Z_n \in M)\times$$

$$\times\, \mathbb{P}_{x_0}(Z_1 = x_1, \dots, Z_{n-1} = x_{n-1}, Z_n \in M))$$

$$= \sum_{\substack{n\geqslant 1\\ x_1,\dots,x_{n-1}\in N}} \frac{P(x_0,x_1)}{1-\beta(x_0)v^2}\cdots\frac{P(x_{n-2},x_{n-1})}{1-\beta(x_{n-2})v^2}\frac{P(x_{n-1},M)}{1-\beta(x_{n-1})v^2}$$

$$= \sum_{n\geq 1} \delta_{x_0}(P_v^{(N)})^{n-1}\varphi\,,$$

for $\varphi(x) = \frac{1}{1-\beta(x)v^2}P(x,M)$. Notice that φ is well-defined since $v^2 < 1/\overline{\beta}$. Moreover it is positive because X is irreducible recurrent, so, for any $x_0 \in N$ there exists a path that leads to M.

If $\rho(P_v^{(N)}) < 1$, then

$$\limsup_{n\to+\infty} \left|\delta_{x_0}(P_v^{(N)})^{n-1}\varphi\right|^{1/n} \leqslant \limsup_{n\to+\infty} \left\|(P_v^{(N)})^n\right\|^{1/n} < 1\,,$$

hence the series is convergent.

If $\rho_v := \rho(P_v^{(N)}) \geqslant 1$, by Perron-Frobenius theorem, there exists a probability measure ν_0 with some positive coefficients such that $\nu_0 P_v^{(N)} = \rho_v\nu_0$, which implies that

$$\mathbb{E}_{\nu_0}(e^{vI_\cdot}) = \nu_0(\varphi)\sum_{n\geqslant 0}\rho_v^{n-1} = +\infty,$$

since φ is positive. □

Remark 3.4. *When X is irreducible in restriction to N (i.e. the matrices $P_v^{(N)}$ are irreducible for any v), then $\mathbb{E}(e^{vI_x}) = +\infty$ for all $x \in N$ as soon as $\rho(P_v^{(N)}) \geqslant 1$. If this it not the case, the previous proposition just ensures that when $\rho(P_v^{(N)}) \geqslant 1$, then $\mathbb{E}(e^{vI_x}) = +\infty$ for some $x \in N$. Moreover, for any $x, x' \in N$ such that $P(x,x')$ is positive then $\mathbb{E}(e^{vI_{x'}}) = +\infty$ implies $\mathbb{E}(e^{vI_x}) = +\infty$.*

We now introduce the sub-process made of the positions of (X,Y) at the successive hitting times of M.

Proposition 3.5. *For any $n \geqslant 0$, let us define*

$$U_n = X_{T_{2n}} \quad \text{and} \quad V_n = Y_{T_{2n}}.$$

The process (U,V) is a Markov chain on $F\times\mathbb{R}$. More precisely,

$$V_{n+1} = M_n(U_n)V_n + Q_n(U_n),$$

where the sequence of random vectors $\big((M_n(x), Q_n(x))_{x\in F}\big)$ *is i.i.d., and independent of* (U_n)*, with law given by*

$$M_n(x) \overset{\mathcal{L}}{=} \exp\left(-\int_0^{T_1} \lambda(X_r^x)\,dr\right)$$
$$Q_n(x) \overset{\mathcal{L}}{=} \int_0^{T_1} \sigma(X_s^x) \exp\left(-\int_s^{T_1} \lambda(X_r^x)\,dr\right) dB_s + \int_{T_1}^{T_2} \sigma(X_s^x)\,dB_s.$$

For any $v < v_c$ *where* $v_c = \sup\left\{v,\ \rho(P_v^{(N)}) < 1\right\}$*, we have*

$$\sup_{n\geqslant 0} \mathbb{E}\left(e^{v|V_n|}\right) < +\infty.$$

Moreover, if $v \geqslant v_c$*, this supremum is infinite.*

Proof. The fact that (U, V) is a recurrent Markov chain is a straightforward application of the Markov property for X.

Let us introduce $\overline{M}_n = \max_{x\in F} M_n(x)$ and $\overline{Q}_n = \max_{x\in F} |Q_n(x)|$. The random variables $((\overline{M}_n, \overline{Q}_n))_{n\geqslant 0}$ are i.i.d. Define the sequence $(\overline{V}_n)_{n\geqslant 0}$ by

$$\overline{V}_0 = |V_0| \quad \text{and} \quad \overline{V}_{n+1} = \overline{M}_n \overline{V}_n + \overline{Q}_n \quad \text{for } n \geqslant 1.$$

The domain of the Laplace transforms of $(\overline{V}_n)_{n\geqslant 0}$ is known thanks to the exhaustive study [1]. Since $\mathbb{P}(\overline{Q}_n = 0) < 1$, $\mathbb{P}(0 < \overline{M}_n < 1) = 1$ and for any $c \in \mathbb{R}$, $\mathbb{P}(\overline{Q}_n + \overline{M}_n c = c) < 1$, [1, Theorem 1.6] ensures in particular that $(\mathbb{E}\exp\left(v\overline{V}_n\right))_n$ is uniformly bounded as soon as the Laplace transform $L_{\overline{Q}}$ of $\overline{Q}$ is finite. At last, for any $v \geqslant 0$,

$$\sup_{x\in F} \mathbb{E}\left(e^{v|Q(x)|}\right) \leqslant \mathbb{E}\left(e^{v\overline{Q}}\right) = \mathbb{E}\left(\sup_{x\in F} e^{v|Q(x)|}\right) \leqslant \sum_{x\in F} \mathbb{E}\left(e^{v|Q(x)|}\right).$$

Thus $L_{\overline{Q}}(v)$ is finite if and only if $\mathbb{E}\big(e^{v|Q(x)|}\big)$ is finite for any $x \in F$. Since $|V_n| \leqslant \overline{V}_n$ for all $n \geqslant 0$, then

$$\sup_{n\geqslant 0} \mathbb{E}\left(e^{v|V_n|}\right) < +\infty$$

as soon as $L_{\overline{Q}}(v)$ is finite.

On the other hand, choose v such that there exists $x_0 \in F$ such that $\mathbb{E}\big(e^{v|Q(x_0)|}\big)$ is infinite. Then, for any $n \geqslant 0$,

$$\begin{aligned}
\mathbb{E}\left(e^{v|V_{n+1}|}\right) &\geqslant \mathbb{E}\left(e^{v|V_{n+1}|}\mathbb{1}_{\{U_n = x_0\}}\right) \\
&\geqslant \mathbb{E}\left(e^{-v|V_n|}e^{v|Q_n(x_0)|}\mathbb{1}_{\{U_n = x_0\}}\right) \\
&\geqslant \mathbb{E}\left(\mathbb{1}_{\{U_n = x_0\}}e^{-v|V_n|}\right)\mathbb{E}\left(e^{v|Q_n(x_0)|}\right).
\end{aligned}$$

The recurrence of U ensures that $\big\{n \geqslant 0,\ \mathbb{E}\big(e^{v|V_n|}\big) = +\infty\big\}$ is infinite.

The last point is to show that $L_{\overline{Q}}(v)$ is finite if and only if $v < v_c$ where v_c is defined by (11). For any $x \in F$, the random variable $Q_n(x)$ is symmetric and its Laplace transform is finite as soon as, for any $\tilde{x} \in N$, the Laplace transform of

$$I_{\tilde{x}} = \int_0^T \sigma(X_s^{\tilde{x}})\, dB_s$$

is finite, which is true for $|v| < v_c$. Indeed, we have for any v

$$\mathbb{E}\Big(e^{vQ_n(x)}|\mathcal{F}_{T_1}\Big) = \exp\left(v\int_0^{T_1}\sigma(X_s^x)\exp\left(-\int_s^{T_1}\lambda(X_r^x)\,dr\right)dB_s\right)\mathbb{E}\big(e^{vI_{\tilde{x}}}\big)_{|\tilde{x}=X_{T_1}}. \quad (14)$$

Proposition 3.3 ensures that, if $|v| < v_c$ then

$$\mathbb{E}\Big(e^{vQ_n(x)}\Big) \leqslant C(v)\mathbb{E}\bigg(\exp\left(\frac{v^2}{2}\int_0^{T_1}\sigma(X_s^x)^2\exp\left(-2\int_s^{T_1}\lambda(X_r^x)\,dr\right)ds\right)\bigg).$$

Denoting $\overline{\sigma}_M = \max_{x\in M}\sigma(x)$ and $\underline{\lambda}_M = \min_{x\in M}\lambda(x)$, one has

$$\mathbb{E}\Big(e^{vQ_n(x)}\Big) \leqslant C(v)\exp\left(\frac{\overline{\sigma}_M^2}{4\underline{\lambda}_M}v^2\right).$$

By the way, $L_{\overline{Q}}$ is finite on $(-\infty, v_c)$.

We assume now that $v \geqslant v_c$. From Proposition 3.3, we know that, in this case, the set $G = \{x \in N,\ \mathbb{E}(e^{vI_x}) = +\infty\}$ is non empty. Using the irreducibility of X and Remark 3.4, one notices that there exists $x_0 \in F$ such that $\mathbb{P}(X_{T_1}^{x_0} \in G) > 0$. From this remark and (14), one has $\mathbb{E}(e^{vQ_n(x_0)}) = +\infty$ which conclude the proof. □

Let us now extend this result to the whole process Y.

Theorem 3.6. *For any $v < v_c$ where $v_c = \sup\Big\{v,\ \rho(P_v^{(N)}) < 1\Big\}$, we have*

$$\sup_{t\geqslant 0}\mathbb{E}\Big(e^{v|Y_t|}\Big) < +\infty.$$

Moreover, if $v \geqslant v_c$, then this supremum is infinite.

Proof. Choose $t > 0$. We have

$$\mathbb{E}\Big(e^{v|Y_t|}\Big) = \sum_{n=0}^{\infty}\mathbb{E}\Big(e^{v|Y_t|}\mathbb{1}_{\{T_{2n}\leqslant t<T_{2n+2}\}}\Big).$$

We write, for $0 \leqslant v < v_c$,

$$\mathbb{E}\Big(e^{v|Y_t|}\mathbb{1}_{\{T_{2n}\leqslant t<T_{2n+2}\}}\Big) = \mathbb{E}\Big(\mathbb{E}\Big(e^{v|Y_t|}\mathbb{1}_{\{T_{2n}\leqslant t<T_{2n+2}\}}|\mathcal{F}_{T_{2n}}\vee\mathcal{F}_t^X\Big)\Big)$$

As in the proof of Proposition 3.5,

$$\mathbb{E}\Big(e^{v|Y_t|}\mathbb{1}_{\{T_{2n}\leqslant t<T_{2n+2}\}}|\mathcal{F}_{T_{2n}}\vee\mathcal{F}_t^X\Big) \leqslant C(v)\exp\left(\frac{\overline{\sigma}_M^2}{4\underline{\lambda}_M}v^2\right)e^{v|Y_{T_{2n}}|}\mathbb{E}\big(\mathbb{1}_{\{T_{2n}\leqslant t<T_{2n+2}\}}|\mathcal{F}_{T_{2n}}\vee\mathcal{F}_t^X\big).$$

By the Markov property applied to X,

$$\mathbb{E}\Big(e^{v|Y_t|}\mathbb{1}_{\{T_{2n}\leqslant t<T_{2n+2}\}}\Big) \leqslant C(v)\exp\left(\frac{\overline{\sigma}_M^2}{4\underline{\lambda}_M}v^2\right)\mathbb{E}\Big(e^{v\left|Y_{T_{2n}}\right|}\Big)\mathbb{P}(T_{2n}\leqslant t<T_{2n+2}).$$

Then, for $0 \leqslant v < v_c$,

$$\mathbb{E}\Big(e^{v|Y_t|}\Big) \leqslant C(v)\exp\left(\frac{\overline{\sigma}_M^2}{4\underline{\lambda}_M}v^2\right)\sup_{n\geqslant 0}\mathbb{E}\Big(e^{v\left|Y_{T_{2n}}\right|}\Big).$$

The generalisation of the case $v \geqslant v_c$ to the whole process is immediate. □

4 Polynomial moments for the switched diffusion

We denote by A_p the matrix $A - p\Lambda$ where Λ is the diagonal matrix with diagonal $(\lambda(1), \dots, \lambda(d))$ and associate to A_p the quantity

$$\eta_p := -\max_{\gamma\in\mathrm{Spec}(A_p)} \mathrm{Re}\ \gamma.$$

The main goal of this section is to establish the equivalence between the positivity of η_p and the existence of a p^{th} moment for the invariant measure ν of Y. We will also give the proof of Point 1 of Theorem 1.5.

Using classical ideas from spectral theory, we first relate η_p with exponential functionals of λ along the trajectories of X:

Proposition 4.1. *For any $p > 0$, there exist $0 < C_1(p) < C_2(p) < +\infty$ such that, for any initial probability measure π on E, any $t > 0$,*

$$C_1(p)e^{-\eta_p t} \leqslant \mathbb{E}_\pi\left(\exp\left(-\int_0^t p\lambda(X_u)\,du\right)\right) \leqslant C_2(p)e^{-\eta_p t}. \tag{15}$$

Proof. Let us define, for any $p > 0$ and $t > 0$, the matrix $A_{(p,t)}$ by

$$A_{(p,t)}(x,\tilde{x}) = \mathbb{E}_x\left(\exp\left(-\int_0^t p\lambda(X_u)\,du\right)\mathbb{1}_{\{X_t=\tilde{x}\}}\right).$$

On the one hand, one remarks that

$$\mathbb{E}_\pi\left(\exp\left(-\int_0^t p\lambda(X_u)\,du\right)\right) = \pi A_{(p,t)}\mathbf{1} \tag{16}$$

where the coordinates of $\mathbf{1}$ are all equal to 1 and π is a probability measure on E seen as a row vector.

On the other hand, a simple application of the Feynman-Kac formula shows that $A_{(p,t)} = e^{tA_p}$. This fact relates the spectra of A_p and $A_{(p,t)}$. In particular, $\rho(A_{(p,t)}) = e^{-\eta_p t}$ and, since all coefficients of $A_{(p,t)}$ are positive, we can apply the Perron-Frobenius Theorem to ensure that $-\eta_p$ is a simple eigenvalue of A_p, all other eigenvalues having a strictly smaller real part. Let $\xi_p < -\eta_p$ be an upper bound for the real parts of these other eigenvalues.

We then define π_p (resp. φ_p) the left (resp. right) eigenvector associated to $-\eta_p$, with positive coefficients, normalized such that $\pi_p(\mathbf{1}) = 1$ (resp. $\pi_p(\varphi_p) = 1$). Applying [5, Thm VII.1.8], we get that for any $t > 0$

$$e^{tA_p} = e^{-\eta_p t}\varphi_p \pi_p + R_p(t),$$

with $\|R_p(t)\|_\infty \leqslant P_p(t)e^{\xi_p t}$, $P_p(t)$ being a polynomial of degree less than d. This gives

$$\pi e^{tA_p}\mathbf{1} = e^{-t\eta_p}(\pi(\varphi_p) + e^{t\eta_p}\pi R_p(t)\mathbf{1})$$

hence

$$e^{-t\eta_p}(\pi(\varphi_p) - P_p(t)e^{t(\eta_p+\xi_p)}) \leqslant \pi e^{tA_p}\mathbf{1} \leqslant e^{-t\eta_p}(\pi(\varphi_p) + P_p(t)e^{t(\eta_p+\xi_p)}).$$

This estimate gives (15) thanks to (16) and to the fact that $P_p(t)e^{t(\eta_p+\xi_p)}$ tends to 0 as t tends to infinity. □

Let us now study the function $p \mapsto \eta_p$.

Proposition 4.2.

1. *The function $p \mapsto \eta_p$ is smooth and concave on $\mathbb{R}_+$. Its derivative at $p = 0$ is equal to*
$$\sum_{x\in E} \lambda(x)\mu(x) > 0,$$
and η_p/p tends to $\underline{\lambda}$ as p goes to infinity.

2. *We have the following dichotomy:*
   - *if $\underline{\lambda} \geqslant 0$, then for all $p > 0$, $\eta_p > 0$,*
   - *if $\underline{\lambda} < 0$, there is $\kappa \in (0, \min\{-a(x)/\lambda(x), \lambda(x) < 0\})$ such that $\eta_p > 0$ for $p < \kappa$ and $\eta_p < 0$ for $p > \kappa$.*

Proof. The smoothness of the functions η_p, π_p and φ_p are classical results of perturbation theory (see for example [9, chapter 2]). Since $\pi_p A_p = -\eta_p \pi_p$, $\pi_p \mathbf{1} = 1$ and $A\mathbf{1} = 0$, one has

$$\eta_p = -\pi_p A_p \mathbf{1} = p\pi_p \Lambda \mathbf{1} = p\sum_{x\in E} \pi_p(x)\lambda(x). \tag{17}$$

Differentiating this relation gives $\eta_p' = \pi_p\Lambda\mathbf{1} + p\pi_p'\Lambda\mathbf{1}$. In particular, $\eta_0' = \mu\Lambda\mathbf{1} = \sum_{x\in E}\mu(x)\lambda(x)$, since $\pi_0 = \mu$.

We turn to the proof of the concavity of η_p. We only have to remark that, for any $t > 0$ and any $x \in E$,

$$p \mapsto M_t^{(x)}(p) = \frac{1}{t}\log \mathbb{E}_x\left(\exp\left(-p\int_0^t \lambda(X_u)\,du\right)\right)$$

is a convex function, as a log-Laplace transform (for example using Hölder's inequality). But (15) implies that $M_t^{(x)}$ converges to $-\eta_p$, hence η_p is concave as a limit of concave functions.

Obviously, one has, for any $t > 0$ and $p > 0$, $M_t^{(x)}(p) \leqslant -p\underline{\lambda}$ and η_p is greater than $p\underline{\lambda}$. On the other hand, denoting by T the first jump time of (X_t), one has

$$\begin{aligned} M_t^{(x)}(p) \geqslant & \frac{1}{t} \log \mathbb{E}_x \left(\exp \left(-p \int_0^t \lambda(X_u)\, du \right) \mathbb{1}_{\{T>t\}} \right) \\ \geqslant & -p\lambda(x) + \frac{1}{t} \log \mathbb{P}_x(T > t) = -p\lambda(x) - a(x). \end{aligned}$$

When t goes to infinity, one gets for any $p > 0$

$$\eta_p \leqslant \min_{x \in E} (a(x) + p\lambda(x)). \tag{18}$$

In particular, η_p/p goes to $\underline{\lambda}$ as p goes to infinity.

The fact that, when $\underline{\lambda} \geqslant 0$, η_p is always positive is clear from (17).

When $\underline{\lambda} < 0$, for p small enough, $\eta_p > 0$ since its derivative at $p = 0$ is positive. But in this case, we can check that $\eta_p < 0$ for p large enough. Equation (18) implies that $\eta_p < 0$ as soon as $p > \min_{x \in E, \lambda(x) < 0} -a(x)/\lambda(x)$. This provides the upper bound for κ.

With the concavity of η_p, these considerations are sufficient to ensure that η_p as a unique zero κ, being positive before and negative after. ☐

Remark 4.3. *The relation $\eta_\kappa = 0$ implies that $(A - \kappa\Lambda)\varphi_\kappa = 0$ which can be rewritten as $M_\kappa \varphi_\kappa = \varphi_\kappa$ (M_κ being the matrix defined in (4)). This ensures that $\rho(M_\kappa) = 1$ since M_κ is non-negative irreducible and φ_κ is positive. By the way our characterization of κ in Theorem 1.5 is equivalent to the one given by de Saporta and Yao in Point 1. of Theorem 1.2.*

It is known from [7, 4] that the invariant measure ν of Y has p^{th} finite moment if and only if $p < \kappa$. Their proof is based on a time discretization of the process (X, Y) together with generic results on the ergodicity of discrete time Markov processes and renewal theory (see [3]). The previous propositions provide a direct and simple characterization of the critical moment of ν.

Proposition 4.4. *For any $p > 0$ such that $\eta_p > 0$ (*i.e. *$p < \kappa$), and any initial measure such that the second marginal has a p^{th} finite moment, one has*

$$\sup_{t \geqslant 0} \mathbb{E}(|Y_t|^p) < +\infty \quad \textit{and} \quad \int |y|^p \, \nu(dy) < +\infty.$$

*On the other hand, for any p such that $\eta_p \leqslant 0$ (*i.e. *$p \geqslant \kappa$) and any initial condition,*

$$\lim_{t \to \infty} \mathbb{E}(|Y_t|^p) = +\infty \quad \textit{and} \quad \int |y|^p \, \nu(dy) = +\infty.$$

Proof. Let us assume that $p \geqslant 2$. If it is not the case, one has to replace the function $y \mapsto |y|^p$ by the $\mathcal{C}^2$ function $y \mapsto \frac{|y|^{p+2}}{1+|y|^2}$. Choose $T > 0$. Itô's formula ensures that

$$d|Y_t|^p = \left(-p\lambda(X_t)|Y_t|^p + \frac{p(p-1)}{2} \sigma(X_t)^2 |Y_t|^{p-2} \right) dt + p\sigma(X_t) Y_t |Y_t|^{p-2} \, dB_t. \tag{19}$$

Let us denote by α_p the function defined on $[0,T]$ by

$$\alpha_p(t) = \mathbb{E}\big(|Y_t|^p|\mathcal{F}_T^X\big).$$

Taking the expectation of (19) conditionnally to X leads to

$$\alpha_p'(t) = -p\lambda(X_t)\alpha_p(t) + \frac{p(p-1)}{2}\sigma^2(X_t)\alpha_{p-2}(t),$$

since B and X are independent. For any $\varepsilon > 0$, there exists c such that

$$\alpha_p'(t) \leqslant (-p\lambda(X_t)+\varepsilon)\alpha_p(t) + c.$$

This implies that

$$\alpha_p(t) \leqslant \alpha_p(0)e^{\int_0^t(-p\lambda(X_r)+\varepsilon)\,dr} + c\int_0^t e^{\int_u^t(-p\lambda(X_r)+\varepsilon)\,dr}\,du.$$

One has to take the expectation and use (15) to get for any $p > 2$ such that $\eta_p > 0$

$$\mathbb{E}(|Y_t|^p) \leqslant C_2(p)\mathbb{E}(|Y_0|^p)e^{(-\eta_p+\varepsilon)t} + c\; C_2(p)\int_0^t e^{-(-\eta_p+\varepsilon)u}\,du.$$

If $\varepsilon < \eta_p$ then $\sup_{t>0}\mathbb{E}(|Y_t|^p)$ is finite.

If $p=\kappa$, one has

$$\alpha_\kappa'(t) = -\kappa\lambda(X_t)\alpha_\kappa(t) + \frac{\kappa(\kappa-1)}{2}\sigma^2(X_t)\alpha_{\kappa-2}(t).$$

Then

$$\begin{aligned}\alpha_\kappa(t) &= \int_0^t e^{-\kappa\int_s^t\lambda(X_u)\,du}\kappa(\kappa-1)\sigma(X_s)^2\alpha_{\kappa-2}(s)\,ds + \mathbb{E}(|Y_0|^\kappa)e^{-\kappa\int_0^t\lambda(X_u)\,du}\\ &\geqslant \kappa(\kappa-1)\underline{\sigma}^2\int_0^t e^{-\kappa\int_s^t\lambda(X_u)\,du}\alpha_{\kappa-2}(s)\,ds.\end{aligned}$$

As a consequence, using Proposition 4.1 and the relation $\eta_\kappa = 0$ (see Proposition 4.2),

$$\begin{aligned}\mathbb{E}(|Y_t|^\kappa) &\geqslant \kappa(\kappa-1)\underline{\sigma}^2\int_0^t \mathbb{E}\Big(\alpha_{\kappa-2}(s)\mathbb{E}\Big(e^{-\kappa\int_s^t\lambda(X_u)\,du}|\mathcal{F}_s^X\Big)\Big)\,ds\\ &\geqslant \kappa(\kappa-1)\underline{\sigma}^2C_1(\kappa)\int_0^t\mathbb{E}\Big(|Y_s|^{\kappa-2}\Big)\,ds.\end{aligned}$$

From the first part of the proof,

$$\lim_{s\to\infty}\mathbb{E}\Big(|Y_s|^{\kappa-2}\Big) = \int |y|^{\kappa-2}\,\nu(dy) > 0.$$

By the way,

$$\lim_{t\to\infty}\mathbb{E}(|Y_t|^\kappa) = +\infty,$$

and the κ^{th} moment of ν is infinite. This is also true for the p^{th} moment for any $p>\kappa$. □

5 Convergence to equilibrium for the switched diffusion

Under the assumption that ν has a finite p^{th} moment, one can establish an exponential convergence of (X,Y) to its invariant measure in terms of mixed total variation (for X) and W_p Wasserstein distance (for Y).

Let us start with the easiest case, assuming that $\mathcal{L}(X_0)=\mathcal{L}(\tilde{X}_0)$.

Proof of Theorem 1.7. Let y and $\tilde{y}$ be two real numbers. We couple two trajectories of (X,Y) starting at (x,y) and $(x,\tilde{y})$ by choosing the same first components and the same Brownian motion to drive Y and $\tilde{Y}$. In other words, we compare $(X_t,Y_t)^{x,y}$ and $(\tilde{X}_t,\tilde{Y}_t)^{x,\tilde{y}}$ where

$$\begin{cases} X_t=\tilde{X}_t, \\ Y_t=y-\displaystyle\int_0^t \lambda(X_u)Y_u\,du+\int_0^t \sigma(X_u)\,dB_u \\ \tilde{Y}_t=\tilde{y}-\displaystyle\int_0^t \lambda(X_u)\tilde{Y}_u\,du+\int_0^t \sigma(X_u)\,dB_u. \end{cases}$$

Then,

$$d\Big(Y_t-\tilde{Y}_t\Big)=-\lambda(X_t)(Y_t-\tilde{Y}_t)\,dt$$

and

$$\Big|Y_t-\tilde{Y}_t\Big|^p=|y-\tilde{y}|^p-\int_0^t p\lambda(X_u)\Big|Y_u-\tilde{Y}_u\Big|^p\,du.$$

As a conclusion, (15) ensures that

$$\mathbb{E}_{(x,y),(x,\tilde{y})}\left(\Big|Y_t-\tilde{Y}_t\Big|^p\right)=\mathbb{E}_x\left(\exp\left(-\int_0^t p\lambda(X_u)\,du\right)\right)|y-\tilde{y}|^p\leqslant C_2(p)e^{-\eta_p t}|y-\tilde{y}|^p.$$

Then, for any coupling Π of $\mathcal{L}(Y_0)$ and $\mathcal{L}(\tilde{Y}_0)$,

$$W_p\Big(\mathcal{L}(Y_t),\mathcal{L}(\tilde{Y}_t)\Big)^p\leqslant C_2(p)e^{-\eta_p t}\int|y-\tilde{y}|^p\,\Pi(d(y,\tilde{y})).$$

Taking the infimum over Π provides the result. ☐

Let us turn to the general case.

Theorem 5.1. *Consider two processes (X,Y) and $(\tilde{X},\tilde{Y})$ with respective initial laws π and $\tilde{\pi}$ two probability measures on $E\times\mathbb{R}$ such that the second marginal has a finite θ^{th} moment with $\theta<\kappa$ (with $\kappa=+\infty$ if $\underline{\lambda}\geqslant 0$). For any $p<\theta$, we have*

$$W_p\Big(\mathcal{L}(Y_t),\mathcal{L}(\tilde{Y}_t)\Big)^p\leqslant C_2(p)(1-p_c)^{1-p/\theta}M_0(\theta)^{p/\theta}\exp\left(-\frac{\gamma\eta_p}{(1-p/\theta)\gamma+\eta_p}t\right)+p_c\overline{W}_p^p e^{-\eta_p t},$$

where

$$p_c=\sum_{x\in E}\mu_0(x)\wedge\tilde{\mu}_0(x)=1-d_{\mathrm{TV}}\Big(\mathcal{L}(X_0),\mathcal{L}(\tilde{X}_0)\Big),$$

$$M_0(\theta)^{p/\theta}=2^p\left(\sup_{t\geqslant 0}\mathbb{E}\Big(|Y_t|^\theta\Big)+\sup_{t\geqslant 0}\mathbb{E}\Big(|\tilde{Y}_t|^\theta\Big)\right)^{p/\theta},$$

$$\overline{W}_p=\max_{x\in E}W_p\Big(\mathcal{L}(Y_0|X_0=x),\mathcal{L}(\tilde{Y}_0|\tilde{X}_0=x)\Big),$$

and γ is such that

$$d_{\mathrm{TV}}(\mathcal{L}(X_t), \mathcal{L}(\tilde{X}_t)) \leqslant e^{-\gamma t} d_{\mathrm{TV}}\Big(\mathcal{L}(X_0), \mathcal{L}(\tilde{X}_0)\Big).$$

Remark 5.2. *This estimate can be improved and simplified if $\underline{\lambda} > 0$. In this case, one can write instead of* (20) *that*

$$\mathbb{E}_{(x,y),(\tilde{x},\tilde{y})}\left(\left|Y_t - \tilde{Y}_t\right|^p \mathbb{1}_{\{T \geqslant \alpha t\}}\right) \leqslant C\mathbb{P}(T \geqslant \alpha t)$$

thanks to the explicit expression (2) *of Y. Since $p\underline{\lambda} \leqslant \eta_p$ this leads to*

$$W_p\Big(\mathcal{L}(Y_t), \mathcal{L}(\tilde{Y}_t)\Big)^p \leqslant C(p)(1-p_c)\exp\left(-\frac{\gamma p\underline{\lambda}}{\gamma + p\underline{\lambda}} t\right) + p_c \overline{W}_p^p e^{-p\underline{\lambda} t}.$$

Proof of Theorem 5.1. We have to consider the case $X_0 \neq \tilde{X}_0$. Given $x, \tilde{x} \in E$ (with $x \neq \tilde{x}$) and $y, \tilde{y} \in \mathbb{R}$, we introduce the three independent processes $(X_t)_{t\geqslant 0}$, $(\overline{X}_t)_{t\geqslant 0}$ and $(B_t)_{t\geqslant 0}$ where the first one is a chain starting at x, the second one is a chain starting at $\tilde{x}$ and the last one is a standard Brownian motion. The process $\tilde{X}$ is defined as follows:

$$\tilde{X}_t = \begin{cases} \overline{X}_t & \text{if } t \leqslant T, \\ X_t & \text{if } t > T, \end{cases}$$

where $T = \inf\left\{t > 0,\ X_t = \overline{X}_t\right\}$. It is well known (since X is a finite irreducible continuous time Markov chain) that there exists $\gamma > 0$ such that

$$\sup_{x,\tilde{x}\in E} \mathbb{P}_{x,\tilde{x}}(T > t) \leqslant e^{-\gamma t}.$$

Let us now define for any $t \geqslant 0$,

$$\begin{aligned} Y_t &= y e^{-\int_0^t \lambda(X_u)\,du} + \int_0^t e^{-\int_u^t \lambda(X_v)\,dv} \sigma(X_u)\,dB_u, \\ \tilde{Y}_t &= \tilde{y} e^{-\int_0^t \lambda(\tilde{X}_u)\,du} + \int_0^t e^{-\int_u^t \lambda(\tilde{X}_v)\,dv} \sigma(\tilde{X}_u)\,dB_u. \end{aligned}$$

Let us denote, for any $p < \kappa$ and $y, \tilde{y} \in \mathbb{R}$,

$$C(p,x,y) = \sup_{t\geqslant 0} \mathbb{E}_{x,y}(|Y_t|^p) \quad \text{and} \quad C(p,x,y,\tilde{x},\tilde{y}) = 2^p(C(p,x,y) + C(p,\tilde{x},\tilde{y})).$$

Let $\alpha \in (0,1)$ and s be the conjugate of θ/p. Theorem 1.7 ensures that

$$\begin{aligned} \mathbb{E}_{(x,y),(\tilde{x},\tilde{y})}\left(\left|Y_t - \tilde{Y}_t\right|^p\right) &= \mathbb{E}_{(x,y),(\tilde{x},\tilde{y})}\left(\left|Y_t - \tilde{Y}_t\right|^p \left(\mathbb{1}_{\{T\geqslant \alpha t\}} + \mathbb{1}_{\{T<\alpha t\}}\right)\right) \\ &\leqslant C(\theta,x,y,\tilde{x},\tilde{y})^{p/\theta} e^{-\gamma\alpha t/s} \\ &\qquad + \mathbb{E}_{(x,y),(\tilde{x},\tilde{y})}\left(\left|Y_T - \tilde{Y}_T\right|^p C_2(p) e^{-\eta_p(t-T)} \mathbb{1}_{\{T<\alpha t\}}\right) \\ &\leqslant C_2(p) C(\theta,x,y,\tilde{x},\tilde{y})^{p/\theta}\left(e^{-\gamma\alpha t/s} + e^{-\eta_p(1-\alpha)t}\right). \end{aligned} \tag{20}$$

Optimizing over α in order to have $\gamma\alpha/s = \eta_p(1-\alpha)$ *i.e.* $\alpha = \frac{s\eta_p}{\gamma+s\eta_p}$ leads to

$$\mathbb{E}_{(x,y),(\tilde{x},\tilde{y})}\Big(\Big|Y_t - \tilde{Y}_t\Big|^p\Big) \leqslant C_2(p)C(\theta,x,y,\tilde{x},\tilde{y})^{p/\theta}\exp\left(-\frac{\gamma\eta_p}{\gamma+s\eta_p}t\right).$$

Let us now turn to the case of general initial conditions. Let π_0 and $\tilde{\pi}_0$ be two probability measures on $E\times\mathbb{R}$ such that the second marginal has a finite θ^{th} moment. Let us start coupling the marginals μ_0 and $\tilde{\mu}_0$ on E. Define the coupling probability p_c

$$p_c = \sum_{x\in E}\mu_0(x)\wedge\tilde{\mu}_0(x),$$

and $D = \{x\in E,\ \mu_0(x) \geqslant \tilde{\mu}_0(x)\}$. We introduce the random variables U, V, W and Z such that for any $x\in E$

$$\begin{aligned}
\mathbb{P}(U=x) &= \frac{\mu_0(x)\wedge\tilde{\mu}_0(x)}{p_c},\\
\mathbb{P}(V=x) &= \frac{\mu_0(x)-\tilde{\mu}_0(x)}{1-p_c}\mathbb{1}_D(x),\\
\mathbb{P}(W=x) &= \frac{\tilde{\mu}_0(x)-\mu_0(x)}{1-p_c}\mathbb{1}_{D^c}(x),
\end{aligned}$$

and $\mathbb{P}(Z=1) = 1-\mathbb{P}(Z=0) = p_c$, Z being independent of (U,V,W). We can now define

$$X_0 = \begin{cases} U & \text{if } Z=1,\\ V & \text{if } Z=0,\end{cases}\qquad \tilde{X}_0 = \begin{cases} U & \text{if } Z=1,\\ W & \text{if } Z=0.\end{cases}$$

We check by a standard computation that the law of X_0 (resp. $\tilde{X}_0$) is μ_0 (resp. $\tilde{\mu}_0$).

Now, for any $x\in E$, let us introduce two random variables Y_0^x and $\tilde{Y}_0^x$, independent of (U,V,W,Z) such that

$$\mathbb{E}\Big(\Big|Y_0^x - \tilde{Y}_0^x\Big|^\theta\Big) = W_\theta\Big(\mathcal{L}(Y_0|X_0=x),\mathcal{L}(\tilde{Y}_0|\tilde{X}_0=x)\Big)^\theta.$$

With this construction $(X_0, Y_0^{X_0})$ has law π_0 and $(\tilde{X}_0,\tilde{Y}_0^{\tilde{X}_0})$ has law $\tilde{\pi}_0$. We consider the processes (X,Y) and $(\tilde{X},\tilde{Y})$ with these initial conditions, the sticky Markov chains and the same Brownian motion. Thanks to the previous computations, we have

$$\begin{aligned}
\mathbb{E}\Big(\Big|Y_t-\tilde{Y}_t\Big|^p\Big) &= \mathbb{E}\Big(\Big|Y_t-\tilde{Y}_t\Big|^p\Big(\mathbb{1}_{\{X_0=\tilde{X}_0\}}+\mathbb{1}_{\{X_0\neq\tilde{X}_0\}}\Big)\Big)\\
&\leqslant \mathbb{E}\Big(\mathbb{1}_{\{X_0=\tilde{X}_0\}}\Big|Y_0^{X_0}-\tilde{Y}_0^{\tilde{X}_0}\Big|^p\Big)e^{-\eta_p t}\\
&\quad + C_2(p)\mathbb{E}\Big(\mathbb{1}_{\{X_0\neq\tilde{X}_0\}}C(\theta,X_0,Y_0^{X_0},\tilde{X}_0,\tilde{Y}_0^{\tilde{X}_0})^{p/\theta}\Big)\exp\left(-\frac{\gamma\eta_p}{\gamma+s\eta_p}t\right).
\end{aligned}$$

On the one hand, we have

$$\begin{aligned}
\mathbb{E}\Big(\mathbb{1}_{\{X_0=\tilde{X}_0\}}\Big|Y_0^{X_0}-\tilde{Y}_0^{\tilde{X}_0}\Big|^p\Big) &= \mathbb{E}\Big(\mathbb{1}_{\{X_0=\tilde{X}_0\}}\mathbb{E}\Big(\Big|Y_0^{X_0}-\tilde{Y}_0^{X_0}\Big|^p|X_0=\tilde{X}_0\Big)\Big)\\
&\leqslant p_c\overline{W}_p^p,
\end{aligned}$$

where $\overline{W}_p = \max_{x\in E} W_p\Big(\mathcal{L}(Y_0|X_0=x), \mathcal{L}(\tilde{Y}_0|\tilde{X}_0=x)\Big)$. On the other hand,

$$\mathbb{E}\Big(\mathbb{1}_{\{X_0\neq\tilde{X}_0\}} C(\theta, X_0, Y_0^{X_0}, \tilde{X}_0, \tilde{Y}_0^{\tilde{X}_0})^{p/\theta}\Big) \leqslant \mathbb{P}(X_0\neq\tilde{X}_0)^{1/s}\mathbb{E}\Big(C(\theta, X_0, Y_0^{X_0}, \tilde{X}_0, \tilde{Y}_0^{\tilde{X}_0})\Big)^{p/\theta}.$$

As a conclusion we get the following bound:

$$W_p\Big(\mathcal{L}(Y_t), \mathcal{L}(\tilde{Y}_t)\Big)^p \leqslant C_2(p)(1-p_c)^{1/s} M_0(\theta)^{p/\theta} \exp\left(-\frac{\gamma\eta_p}{\gamma+s\eta_p}t\right) + p_c^{1/s}\overline{W}_\theta^{p/\theta} e^{-\eta_p t},$$

where

$$M_0(\theta)^{p/\theta} = 2^p\Big(\mathbb{E}(C(\theta, X_0, Y_0)) + \mathbb{E}\Big(C(\theta, \tilde{X}_0, \tilde{Y}_0)\Big)\Big)^{p/\theta}.$$

□

References

[1] G. Alsmeyer, A. Iksanov, and U. Rösler, *On distributional properties of perpetuities*, J. Theoret. Probab. **22** (2009), no. 3, 666–682. MR MR2530108 1, 3.2

[2] G. K. Basak, A. Bisi, and M. K. Ghosh, *Stability of a random diffusion with linear drift*, J. Math. Anal. Appl. **202** (1996), no. 2, 604–622. MR MR1406250 (97g:60091) 1

[3] B. de Saporta, *Tail of the stationary solution of the stochastic equation $Y_{n+1} = a_nY_n + b_n$ with Markovian coefficients*, Stochastic Process. Appl. **115** (2005), no. 12, 1954–1978. MR MR2178503 (2006g:60129) 4

[4] B. de Saporta and J.-F. Yao, *Tail of a linear diffusion with Markov switching*, Ann. Appl. Probab. **15** (2005), no. 1B, 992–1018. MR MR2114998 (2005k:60257) 1, 1, 1.2, 1.3, 4

[5] N. Dunford and J. T. Schwartz, *Linear operators. Part I*, Wiley Classics Library, John Wiley & Sons Inc., New York, 1988, General theory, With the assistance of William G. Bade and Robert G. Bartle, Reprint of the 1958 original, A Wiley-Interscience Publication. MR MR1009162 (90g:47001a) 4

[6] C. M. Goldie and R. Grübel, *Perpetuities with thin tails*, Adv. in Appl. Probab. **28** (1996), no. 2, 463–480. MR MR1387886 (97f:60124) 1

[7] X. Guyon, S. Iovleff, and J.-F. Yao, *Linear diffusion with stationary switching regime*, ESAIM Probab. Stat. **8** (2004), 25–35 (electronic). MR MR2085603 (2005h:60244) 1, 1, 1, 1.3, 4

[8] P. Hitsczenko and J. Wesołowski, *Perpetuities with thin tails revisited*, Ann. Appl. Probab. **19** (2009), no. 6, 2080–2101. 1

[9] T. Kato, *Perturbation theory for linear operators*, Classics in Mathematics, Springer-Verlag, Berlin, 1995, Reprint of the 1980 edition. MR MR1335452 (96a:47025) 4

[10] J.R. Norris, *Markov chains*, Cambridge Series in Statistical and Probabilistic Mathematics, 1997. 1

[11] C. Villani, *Topics in optimal transportation*, Graduate Studies in Mathematics, vol. 58, American Mathematical Society, Providence, RI, 2003. MR MR1964483 (2004e:90003) 1

`Compiled December 17, 2009.`

Jean-Baptiste Bardet e-mail: `jean-baptiste.bardet(AT)univ-rouen.fr`

UMR 6085 CNRS Laboratoire de Mathématiques Raphaël Salem (LMRS)
Université de Rouen, Avenue de l'Université, BP 12, F-76801 Saint Etienne du Rouvray

Hélène Guérin, e-mail: `helene.guerin(AT)univ-rennes1.fr`

UMR 6625 CNRS Institut de Recherche Mathématique de Rennes (IRMAR)
Université de Rennes I, Campus de Beaulieu, F-35042 Rennes Cedex, France.

Florent Malrieu, corresponding author, e-mail: `florent.malrieu(AT)univ-rennes1.fr`

UMR 6625 CNRS Institut de Recherche Mathématique de Rennes (IRMAR)
Université de Rennes I, Campus de Beaulieu, F-35042 Rennes Cedex, France.